AD ________________

Award Number: DAMD17-03-2-0015

TITLE: REMIS: A Tool to Facilitate Innovation, Standardization and Validation in the Creation of Ergonomic Surgical Simulators

PRINCIPAL INVESTIGATOR: W. Brent Seales, Ph.D.

CONTRACTING ORGANIZATION: University of Kentucky Research Foundation
Lexington, KY 40506-0057

REPORT DATE: February 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
February 2004

3. REPORT TYPE AND DATES COVERED
Annual (6 Jan 2003 - 5 Jan 2004)

4. TITLE AND SUBTITLE
REMIS: A Tool to Facilitate Innovation, Standardization and Validation in the Creation of Ergonomic Surgical Simulators

5. FUNDING NUMBERS
DAMD17-03-2-0015

6. AUTHOR(S)
W. Brent Seales, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Kentucky Research Foundation
Lexington, KY 40506-0057

E-Mail: seales@uky.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

The goal of this work is to develop and test new technologies that will break down the barriers that block more surgeons from attaining and continuing to practice (without injury or pain) high levels of skill in minimally invasive surgery (MIS). This project will develop new technology by concentrating on three major research thrusts: **Smart Image:** the project will develop and evaluate new approaches for extracting, fusing, and presenting information cues from imagery and other data sources; **Configurable Display:** the project will develop new approaches for presenting existing data (video, CT data) and extracted cues (3D reconstruction, haptic cues, etc.) to the user within a flexible, configurable display environment; **Ergonomic Assessment:** the project will use existing technology and build new techniques as needed to acquire crucial ergonomic data relative to key factors of patient position, technology configuration, and instrument design.

14. SUBJECT TERMS
Minimally Invasive Surgery, Visualization, Ergonomic Assessment, Simulation, Modeling, Calibration, Smart Image, Acquisition

15. NUMBER OF PAGES
18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

INTRODUCTION

Information cues available in laparoscopy and other forms of minimally invasive surgery are impoverished relative to cues available in open surgery. Acquiring surgical skill in such an environment is extremely challenging. Even after mastery, continued practice can lead to problems for the surgeon as indicated by frequent incidence of pain and injury associated with laparoscopy. The long-term impact on the surgeon performing these procedures is largely unknown.

The goal of this work is to develop and test new technologies that will break down the barriers that block more surgeons from attaining and continuing to practice (without injury or pain) high levels of skill in MIS. This project will develop new technology by concentrating on three major research thrusts:

a. **Smart Image**: the project will develop and evaluate new approaches for extracting, fusing, and presenting information cues from imagery and other data sources.

b. **Configurable Display**: the project will develop new approaches for presenting existing data (video, CT data) and extracted cues (3D reconstruction, haptic cues, etc.) to the user within a flexible, configurable display environment

c. **Ergonomic Assessment**: the project will use existing technology and build new techniques as needed to acquire crucial ergonomic data relative to key factors of patient position, technology configuration, and instrument design.

BODY

With the approval of the Grants Officer, this project has undergone substantial personnel changes in its first year. These changes have shifted the timeline for completing project milestones outlined in the approved "Statement of Work." The new timeline for completing proposed work is reflected in this annual report.

Specifically, the project has undergone the following changes in the past year:

- Change of PI (from Adrian Park, M.D. to W. Brent Seales, Ph.D.)
- Modification in Statement of Work
- Establishment of two new subcontracts

The Principal Investigator was changed in response to the change in affiliation of the original Principal Investigator, Adrian Park, M.D., who left the University of Kentucky in September 2003 in order to accept a position as Head of General Surgery at the University of Maryland Medical Center in Baltimore, MD. The central role of Information Technology in this work made it possible for W. Brent Seales, a collaborator with the original Principal Investigator, to obtain approval from the Grants Officer to assume the role of Principal Investigator. This change was approved and executed in

conjunction with the establishment of a subcontract to the University of Maryland to continue progress on the clinical portion of the project through the work of Adrian Park, M.D.

The administrative re-organization of the project necessitated a pause in project effort, limiting progress toward project milestones for a period of time. The Grants Officer has been cognizant and supportive at all times during this period of reorganization. In this report we show two months of progress from the original 13-month schedule. The shift is the result of protracted negotiations leading to the relocation of Adrian Park, M.D. to Maryland, together with the administrative overhead of the change of principal investigator and the time cost of moving and re-establishing the development laboratory in a new location.

In the following we report progress relative to Information Technology milestones (a) through (e) from the approved "Statement of Work:"

Primary Milestones: IT

a. Hire and train development personnel. (months 1-4)
b. Establish software development environment. (hardware, software, standards) (months 2-4)
c. Establish MIS test-bed. (collect instruments, interfaces) (months 4-6)
d. Gather product requirements. (months 1-4)
e. Analyze product requirements. (months 2-5)
f. Hold a multi-site project workshop meeting. (month 10)
g. Define a scalable real-time architecture for multi-input, multi-transformation, multi-display MIS support. (months 4-6)
h. Identify re-usable components from public domain and outside researchers. (months 6-8)
i. Implement first iteration of architectural framework. (months 7-12)
j. Implement flexible display back-end for unprocessed probe camera data. (months 5-10)

We have hired two full-time developers, each with M.S. degrees in computer science specialized in the area of computer vision. These developers have established a software and hardware development environment according to current industry standards under the direct supervision of Duncan Clarke, Ph.D. (project technical lead and software engineering expert). The two full-time developers have systematically gathered requirements and have worked with project leaders to configure the working environment for an MIS test-bed. This process has been accelerated through the help of Adrian Park, M.D. and his associates at Stryker Endoscopy, who have provided MIS instruments and consultation as needed.

The initial progress toward completion of milestones (a)-(e) is reflected by the successful

publication of an abstract at MMVR 2004, included in the appendix of this report.

We are able to report progress relative to Ergonomic Assessment milestones (a) through (c) from the approved "Statement of Work:"

Primary Milestones: Ergonomic Assessment

a. Obtain UMD IRB/IACUC approval. (months 0-2)
b. Hire and train development personnel. (months 1-4)
c. Establish ergonomic assessment environment (space, hardware, software, assessment methodologies). (months 1-4)
d. Organize human subjects for baseline tasks. (months 2-4)
e. Analyze assessment hardware/software system requirements. (months 2-5)
f. Perform baseline assessment study using human subjects completing tasks in training environment. (months 4-6)
g. Define technical requirements for improved assessment technology. (months 4-6)
h. Process and interpret baseline study using trainer boxes and assessment environment. (months 6-8)
i. Hold a multi-site project workshop meeting. (month 10)
j. Conduct follow-on study with humans completing laparoscopic baseline tasks on animals. (months 8-10)

Ergonomic assessment milestones are to be completed at the University of Maryland under a subcontract requiring new IRB approval. (This is beyond the original approval which was secured at the University of Kentucky.) We have obtained University of Maryland IRB approval and are completing the IACUC process now. We are in the first phase of hiring personnel (b) and securing space at University of Maryland (c) for hardware setup and evaluation.

We do not anticipate substantial deviation from the approved "Statement of Work" other than the stated time shift due to project reorganization.

KEY RESEARCH ACCOMPLISHMENTS

At this point our key research accomplishments primarily involve the establishment of a stable first-rate working environment from which to perform the work:

- Recruitment and training of two full-time software developers with M.S. degrees in Computer Science and detailed domain knowledge in computer vision.
- Set up of commercial-grade software/hardware development environment.
- Renovation and configuration of MIS experimental/simulation test bed.
- Initial set of requirements specifications developed for upcoming implementation and development cycle.
- Publication and presentation of abstract in MMVR 2004.
- Securing IRB approval at University of Maryland.

REPORTABLE OUTCOMES

- Manuscript: MMVR 2004.
- Manuscript: Seminars in Laparoscopic Surgery.
- Electronic project software archive conforming to commercial software principles.
- Electronic archive of project progress and white-paper development.

CONCLUSIONS

Work is currently on schedule relative to our current "Statement of Work" and the revised schedule that commenced in early December with receipt of final approval and authorization to resume work. We anticipate successful completion of proposed work without substantial deviations or a reduction in scope. We will produce an intermediate report of "so what" conclusions as soon as more substantive progress has been made toward project deliverables.

REFERENCES

- Seales, W.B. and Caban, J. "Reconstruction and Enhancement in Monocular Laparoscopic Imagery", Medicine Meets Virtual Reality (MMVR 2004), Newport Beach, California.
- Seales, W.B. and Caban, J. "Visualization Trends: Applications in Laparoscopy", Seminars in Laparoscopic Surgery. October 2003.

APPENDICES

We include copies of two manuscripts that have appeared as a result of this project:

- Seales, W.B. and Caban, J. "Reconstruction and Enhancement in Monocular Laparoscopic Imagery", Medicine Meets Virtual Reality (MMVR 2004), Newport Beach, California.
- Seales, W.B. and Caban, J. "Visualization Trends: Applications in Laparoscopy", Seminars in Laparoscopic Surgery. October 2003.

Reconstruction and Enhancement in Monocular Laparoscopic Imagery

Jesus J. Caban *
W. Brent Seales
Computer Science Department
University of Kentucky

Abstract

Constrained minimally-invasive surgical environments create a number of challenges for the surgeon and for automated tools designed to aid in the performance and analysis of complex procedures. The 3D reconstruction of the operative field opens up a number of possibilities for immersive presentation, automated analysis, and post-operative evaluation of surgical procedures.

This paper presents a method for estimating complete 3D information about scope and instrument positioning from monocular imagery. These measurements can be used as the basis for deriving and presenting additional cues during procedures, and can also be used for post-procedure analysis such as objective estimates of high-level performance measures like economy of motion and ergonomic metrics.

1 Introduction

Minimally-invasive surgery (MIS) provides a number of benefits to the patient, including lower risk of infection and swifter recovery times. Minimal invasion is accomplished primarily through the use of a camera (endoscope) and other surgical instruments inserted into the abdominal cavity through small "keyhole" incisions. The surgeon navigates the process by viewing imagery on an external display. By moving the endoscope and instruments, complex surgical tasks can be efficiently accomplished.

In this work we present a method to extract and make explicit information that is implicitly confounded in the imagery. Such information, though valuable as a direct cue, is usually subtle, especially in monocular imagery. Extracting an explicit representation can provide a ready cue or an analytical tool that otherwise would remain subtle and far less useful. In particular, we concentrate on the problem of recovering the 3D position and orientation of instruments within the endoscope's view, as well as the distance of these instruments from the scope, from each other, and from the anatomy.

Providing 3D information is crucial in addressing one of the primary technical and visual obstacles in conducting MIS procedures, which is the lack of an explicit depth cue. Experts become very good at understanding 3D relationships from monocular imagery, which does not make depth explicit but does contain a number of subtle depth cues, such as perspective distortion and scale, expert knowledge of instrument size, shape and relative positioning, and narrow depth of field which provides a focus cue.

We believe that the ability to extract precise depth measurements, including the position and orientation of instruments, scopes and anatomy, can substantially enhance laparoscopic environments of the future. In particular, we envision two immediate uses when depth information can be made explicit for tracked instruments and anatomy: enhanced visualization for the surgical team, and objective performance measures given video of training and simulation cases.

In the case of enhanced visualization, depth cues extracted from monocular imagery can be used to provide alternative views of instrument position, for example. Top views, closing distance, velocity and accelerations between two instruments, and real metric measurements within the operative field to give an accurate sense of the scale of what is potentially a very small operating space are all possible when depth can be extracted from the imagery.

With respect to performance measures, we can formulate motion *signatures* of instruments as a function of time, incorporating distance, orientation, and the derivatives of these values over time. We believe such measures can give a precision to the problem of analyzing the performance of a task completed within a simulation environment or a training box.

We believe that both enhanced visualization and objective performance metrics are valuable. They rely on the extraction of metric 3D information. Stereo scope systems, such as the *da Vinci* System [1], use a stereo endoscope to

*(jesus@netlab.uky.edu) Computer Science Department, University of Kentucky, 2nd Floor Hardymon Building, Lexington KY 40506.

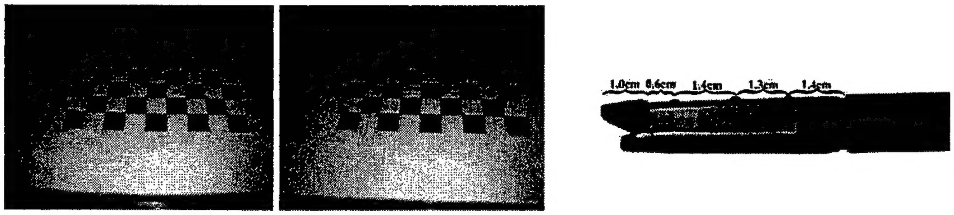

Figure 1: (left) Original image obtained from a 35^o endoscope. (center) Image after undistortion. (right) Stapler instrument with identifiable points and known measurements.

make depth explicit for a viewer using a corresponding stereo viewing system. These systems can also benefit from enhancements and metrics based on depth, which is made even more accessible through the introduction of the stereo scope. One of the crucial contributions of this work is that even when a stereoscopic scope is *not* available, we can obtain certain metric depth measurements. This allows us to apply enhancements and metrics to environments such as trainer boxes, archival procedure video, and current operating areas where it is not practical or even desirable to use specialized hardware like the Da Vinci stereo scope/display.

In the following sections we show how to model and calibrate the camera, and we outline the cue extraction process. We follow this with results from real video that indicate the value of metrics based on extracted cues as we have formulated them. Related work for visual tracking of laparoscopic surgery [3, 2] shows that there continues to be a need for new enhancement techniques, metrics for analysis, and exploration of how best to present extracted cues to the surgical team.

2 The Calibrated Endoscope

In order to formally model the geometry of the endoscope, we assume that the imaging system can be modeled as a pinhole system (i.e., perspective projection). Using this camera model, we apply computer vision methods and algorithms in order to calculate its characteristic geometry and distortion parameters.

First, we compute endoscope parameters and characteristics through a calibration process. According to the pinhole model, the relationship between a 3D point M and its 2D image projection m is given by $sm \approx A[Rt]M$ with

$$A = \begin{pmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

where s is an arbitrary scale factor, (R, t) are the extrinsic parameters, and A is a representation of the intrinsic parameters. Extrinsic parameters locate the camera in a 3D world coordinate frame; intrinsic parameters describe internal camera features such as the pixel coordinates of the image center and the focal length. More specifically, A is a 3x3 matrix that relates the pixel coordinate system to the world coordinate system. Contained in A are parameters α_u and α_v, which together represent the focal length and express the total magnification of the imaging system that results from both optics and image sampling. Also contained in A are the parameters u_0 and v_0, which represent the pixel coordinates of the image center.

As a result of the radial curvature of camera lens elements, there is no real lens system that can produce perfect pinhole images. In the case of endoscopes, different viewing angle scopes, which enlarge the field of view, and scopes with wide-angle lenses, cause significant distortion. These distortions can be removed by calculating distortion parameters through optical calibration. After a camera has been calibrated, it is possible to use the camera parameters to resample any image taken by that camera so that its lens distortion is removed from the image. Figure 1 (left and center) shows lens distortion in an image captured from a 35^o endoscope compared with a distortion-free image generated after calibration.

The key assumption that enables depth reconstruction of instruments visible in monocular sequences is knowledge of the shape, size and the metric measurements of visually identifiable fiducials on the instrument. Based on this information, it is possible to track features in imagery and recover the 3D position of each tracked point. From these points, with *a priori* information about the instrument, it is possible to compute the 3D position and orientation of the tip of the instrument. Figure 1 (right) shows a stapler instrument with identifiable marks and known distances between each of the points.

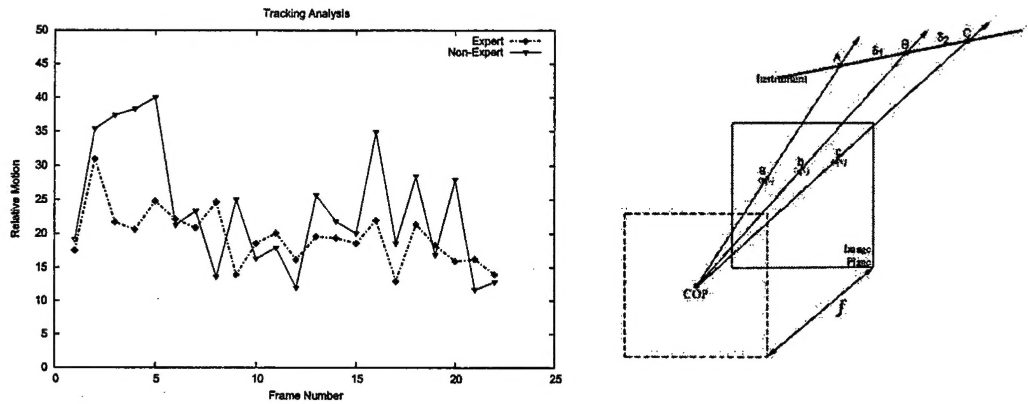

Figure 2: (left) The widely-varying curve shows novice performance; smoother motion is achievable by the expert. (right) The linear constraint leads to a depth solution for known points in monocular imagery.

3 Reconstruction Method

The shape of MIS instruments is almost always linear so that the instrument can be smoothly inserted into ports and manipulated through small incisions. We exploit this fact and as a result simplify the problem of tracking and estimating 3D points that lie on the instrument. By selecting $R = I$ and $t = 0$ in Equation 1, we can compute that $A = z_A A^{-1}a$, $B = z_B A^{-1}b$, and $C = z_C A^{-1}c$, where z_A, z_B and z_C are the unknown depths of the A, B, C points. These equations and constraints lead to a solution [4] for unknown depth wherever the instrument appears in the imagery. These depth values are derived based on the assumption that the instruments are linear, the camera is calibrated (i.e., the projection matrix is available from the off-line calibration process), and the distances between points on the instrument are known *a priori*.

4 Results and Conclusion

We have calibrated various endoscopes (e.g., 0^o and 35^o lenses). After calibration, we used recovered lens distortion estimates to remove lens distortion from images. Using the distortion-corrected images, we tracked the shaft of a stapler instrument (Fig. 2 (right)) in order to recover estimates of the 3D coordinates of points on the instrument. We have found these methods to be very promising as a way to recover 3D cues from monocular data.

As an example, consider the graph in Fig. 2 (left). The two curves on this graph show 3D motion estimates for the stapler instrument over a set of frames. The value plotted as the height of the curve for each frame value is the 3D position of the instrument measured relative to a fixed point. The curve that corresponds to the expert performing the stapling action shows much less relative-motion variation than the curve corresponding to the novice. In this case, economy of motion over a set of frames, evaluated in 3D to capture movement toward and away from the camera, shows how an expert handles the instrument in a way that is measurably and objectively different from the novice.

A number of interesting cues that exist in a subtle, implicit way in monocular laparoscopic imagery can be recovered automatically, enhanced and made explicit. This work shows an example based on the recovery of 3D information relative to the frame of the camera system from points detected on visible instruments as well as points in the operative field tracked over time. These methods do not assume a redesigned camera system, and therefore can be of value for experts using and training with monocular systems today.

References

[1] SunnyBrook Technologies Inc. In *http://www.intuitivesurgical.com.*

[2] R.Raskar et. all H. Fuchs, M. Livingston. Augmented reality visualization for laparoscopic surgery. In *MICCAI 1998*, pages 934–943, Oct 11-13, 1998.

[3] Xiaoli Zhang and Shahram Payandeh. Application of visual tracking for robot-assisted laparoscopic surgery. In *Journal of Robotics Systems*, pages 315–328, Vol. 19, No. 7, 2002.

[4] Z. Zhang. Camera Calibration with One-Dimensional Objects. In *Microsoft Technical Report MSR-TR-2001-120*, August 2002.

Visualization Trends: Applications in Laparoscopy

W. Brent Seales, PhD and Jesus Caban, BS

Recent advances in visualization technology are being driven by two important trends: (1) continued increases in speed and function of hardware devices and (2) increasingly parallel, distributed, cooperative systems. The incorporation of fast, powerful devices into cooperative systems enables a complex interplay of sensors, displays, and computational components that can create a seamless, perceptually rich and flexible environment. Although these trends have fueled a number of advances in visualization research, the unique requirements of laparoscopy make direct, effective use of visualization technology as it is applied in other contexts extremely challenging. This article discusses promising new capabilities in visualization technology. The costs and tradeoffs create new challenges, which are addressed in some visualization applications, but must be carefully assessed in the context of the laparoscopic environment. Incorporating new visualization technology in a way that captures its benefits and meets stringent laparoscopic requirements will very likely precipitate an enormous surge forward in the capabilities of the surgical team and in the quality of patient care.

The laparoscopic environment is intentionally constrained in order to benefit the patient. Minimizing the damage in establishing access to the operative field greatly improves recovery times and lessens the risk of postoperative complications. Although these constraints benefit the patient, they create a physical and informational bottleneck that spawns a number of serious technical challenges. For the patient, "less is more"—less invasive access is more beneficial. For the surgical team, however, "less is more" means that less invasive access leads to procedures that are more challenging, require more skill, more training, and reliance on more technologically sophisticated tools. This article focuses on how current trends in visualization technology may be applied in laparoscopy to put more and more of the burden on the technology, enabling the surgeon to give the patient more benefit with less risk.

Substantial technical development has established an environment that is both minimally invasive and procedurally viable. Although this environment is now widely used for a variety of procedures, the construction of its interface lags far behind current visualization technology. Newly developed visualization technology moves beyond the constraints of the windows, icons, menus, and pointing devices (WIMP) interface and its associated computer engine (the tightly-coupled, single CPU computer).[1,2] In application to laparoscopy, this can create potential advances by breaking barriers that may seem now to be inherent and unavoidable. For example, it may seem that the best laparoscopic view of the operative field must inherently be a two-dimensional (2D) image sequence, captured by a camera and delivered to the surgeon on a 2D computer monitor. As other work has shown, however, this 2D assumption can and is being challenged.[3]

Similarly, it may seem that preoperative data, such as computed tomographic (CT) scans and magnetic resonance imaging (MRI) cannot play a role in the constrained, minimally invasive surgical (MIS) environment because of the disparities and barriers involved: preoperative data may not be available in the operating room, 3D fixed-scan data cannot be adequately aligned with 2D live-camera views, and the surgical team cannot effectively control how to optimize, navigate, and control the many possible ways to fuse different data. The question is "can new visualization technologies continue to expand the limits?"

Of interest is the similarity between the operating room as a whole, with its sterile boundary that divides and separates it from the outside context, and the minimally invasive environment that has its own set of boundaries and constraints. Set within the larger context of the operating room, minimally invasive procedures establish a second boundary that divides and separates the surgeon in the operating room from the operative field. The same challenges in surmounting the required barriers of the operating room apply to the embedded environment of the MIS operative field. It is very likely that the same technologies will be the basis for solutions.

What elements of visualization environments are emerging with current trends? The crucial components are acquisition, modeling and processing, and

From the University of Kentucky, Computer Science Department, Lexington, Kentucky.

Address reprint requests to W. Brent Seales, PhD, Associate Professor, Computer Science Department, University of Kentucky, 773 Anderson Hall, Lexington, KY 40506; E-mail: seales@netlab.uky.edu

display. In the current laparoscopic environment, the camera is a sensor for image acquisition, the modeling and processing step is embedded within the camera hardware, and a cathode ray tube (CRT) is the display device. As shown in Figure 1, the scope is tightly coupled with the display, and the surgical team must integrate information between and among devices as well as the patient. It can require substantial effort for the surgeon to integrate and correlate external views of the patient (Figure 1, right) with simultaneous internal views visible on the display.

Within the operating room, there is very little notion of computation for the purpose of cue detection, enhancement, modeling, and integration—all devices function separately. It is the surgical team's responsibility to interpret the imagery and continuously evaluate the procedure by integrating and understanding information from a parallel set of completely separate devices.[4]

Current visualization environments address this problem by going beyond the WIMP metaphor and grappling with the issues that are required to support immersion, interaction, scalability, and data integration from diverse sources. The requirements of the application are crucial, since they dictate what can be traded in order to develop a workable solution for the problem at hand. The point is to identify trends in visualization, including tradeoffs and task-based requirements, and consider how to appropriately exploit the technology for improvements in laparoscopy.

It is important to note the clear distinction between the operating room environment and practice-oriented training and simulation environments. Training environments continue to advance at a rapid rate and provide a safe test bed for developing new ideas.[5,6] The requirements are much higher in the operating room, however. The goal is to deploy new, revolutionary technology in the operating room and to develop a simulation environment that so closely resembles surgical procedures that it is indistinguishable. Toward that goal, simulation and training tests, especially for visualization technology, are an extremely useful strategy for assessing the potential of a new technology. Success in simulation is a precursor to success at bringing new technology to the operating room.

It is also important to observe that visualization is not the same thing as virtual reality. In fact, equating these two may unnecessarily bias solutions and approaches toward aspects of visualization that are highlighted in virtual reality systems, such as "spatial immersion." What is necessary is a careful look at the requirements and the structure of the current laparoscopic environment, with the continued goal of improving it. The cognitive burden is already great within the walls of the operating room; it is intensified by the added constraint of minimal invasion. Successful visualization technology for the laparoscopist must lessen that burden and yet increase the ability to form and execute a complex, successful procedure under heavy constraints.

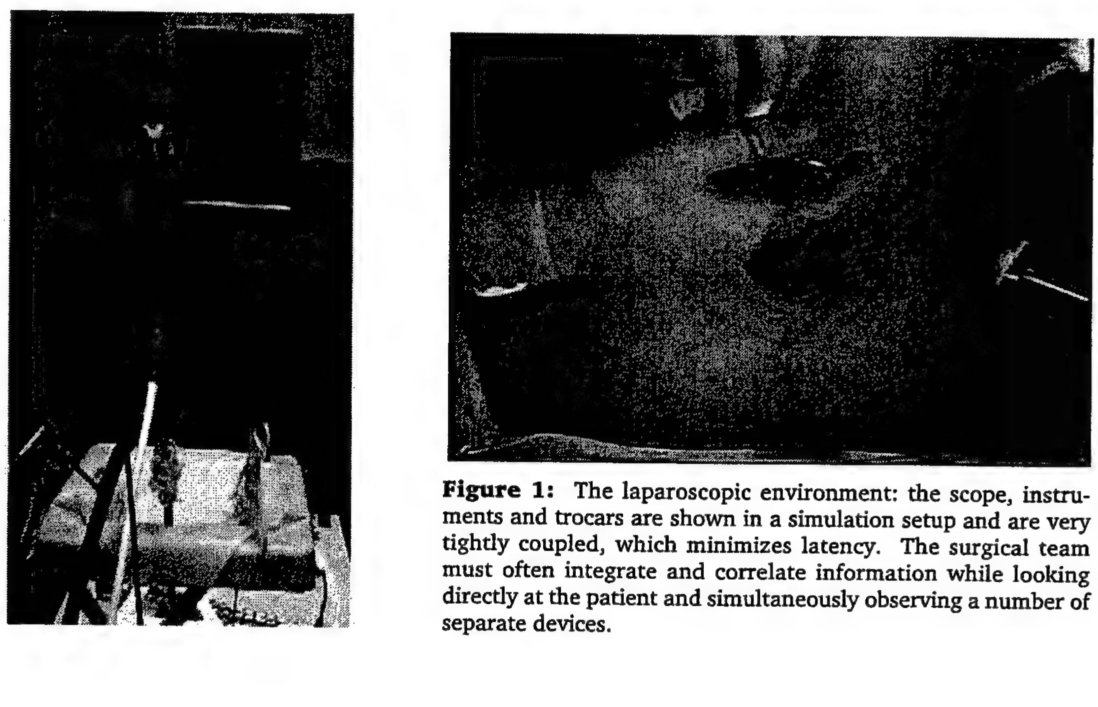

Figure 1: The laparoscopic environment: the scope, instruments and trocars are shown in a simulation setup and are very tightly coupled, which minimizes latency. The surgical team must often integrate and correlate information while looking directly at the patient and simultaneously observing a number of separate devices.

Visualization Technology

Visualization is the human activity of exploring, interpreting, and understanding complex data. The computational aspect of visualization is a transformation of symbols: turning numbers and raw data into geometry, which enables the human observer to interpret complex simulations and phenomena.[7] The data that form the basis of this transformation pipeline must be acquired, processed, and displayed. Technology is critical at each of these points.

Accepted laparoscopic surgical environments rely on imagery acquired with a camera system that illuminates and images the region of interest.[8,9] After acquisition, the imagery is processed and manipulated in order to improve its presentation to the viewer. The imagery is then presented on a display system, most often a CRT or flat panel display, which is positioned within easy view of the surgical team.

Advances in visualization environments derive from improvements in data acquisition, modeling and processing, and display, and in the overall architecture to support the interplay between these elements. The development of new types of acquisition technology must be supported by expanded algorithms and architectures for manipulating and structuring the acquired data as well as new display technology that efficiently communicates this structure to the observer.[10] As a system, the information bandwidth is increasing at the sensor and at the display, supported by more processing power in between.

Sensors and Acquisition

Both sensors and algorithms (simulations) can produce data for visualization. Sensors function at the point of contact with a physical phenomenon in order to sample its essential characteristics. Simulation data are obtained from an algorithm. Often a visualization environment, such as weather simulation, teleimmersion, telecollaboration, and distance learning, combines real and simulated data.

The continued trend of exponential speedup in computing[11,12] and device miniaturization directly affects both real data acquisition and the production of simulation data. Smaller devices equipped with more processing power can detect and deliver better data, such as higher-resolution cameras-on-a-chip and 3D sensors.[13,14] Miniaturization is certainly a welcome trend for laparoscopy, where smaller sensors support the goal of minimal invasion. Every indication is that devices will continue to get smaller and more powerful. For example, the Programmable Digital Camera project[15] investigates algorithms, circuits, and different architectures for portable, small, flexible, and high-resolution digital camera sensors.

As individual sensors increase in resolution and decrease in size, an equally profound trend is that many more sensors are being used cooperatively within a single system. Rather than relying on one, super-high-resolution camera, for example, systems use a number of distributed, lower-resolution cameras. The Office of the Future environment uses networks of cameras and microphones to capture and then remotely render images and audio.[16] That visualization environment is constructed as a way to promote immersive collaboration, and it does so through cooperative, distributed sensors (cameras and microphones) in order to create a sense of presence between physically separate sites.

The use of many sensors increases the computational load and creates the problem of data fusion. Systems have found solutions by coupling the distributed sensing network with a distributed computing architecture that can process data in parallel and fuse data through distributed algorithms that communicate via high-speed network connectivity. As shown in the architectural diagram of Figure 2, clusters of machines connected via a high-speed network manage tasks such as device integration and control, complex simulation processing, and display. This departure from the "single device" architecture (Figure 2, left) supports scalability, parallelism, and information fusion.

Large-scale visualization environments almost all rely on clusters of machines to perform rendering and to drive large-scale simulations.[17-19] Within surgical simulation, the Virtual Reality Based Laparoscopic Surgery Simulator project[20] uses an architecture based on a parallel processing engine to support the communications, computation, and processing needed for laparoscopic applications. Similarly, the daVinci system[21] implements a tightly coupled set of processes to control servos and manage a user interface.

Distributed miniature sensors that are driven by a distributed computing environment are bound by the need to communicate in order to organize and structure the disparate data being collected. Tethered connectivity for the purpose of communication and power may be acceptable in visualization environments but could be a large barrier in surgical environments. The Metaverse

project, for example, uses a communication network to support a grid of cameras that cooperatively senses the state of the display area and makes geometric and photometric corrections when required.[22] In that application, it is of little consequence that the cameras are tethered to provide power and network, since the display area can readily support that infrastructure. However, in surgical environments, where the area of interest is within the body, wireless transmission and associated protocols can be very important. The Massachusetts Institute of Technology's Wireless Microsensor System project[23] emphasizes the study of protocols, networks, and designs for flexible, wireless, and rapidly configurable arrays of sensors.

Increasingly powerful, small, distributed, untethered sensors have brought new capabilities and unprecedented flexibility to visualization environments. Their use also brings a new set of problems that must be solved, including the need for a computational framework capable of fusing data from the sensor network together with other data relevant to the task. The solution adopted in most visualization environments is the distributed "cluster" architecture with high-speed communication connections for data transfer and access.

Modeling and Processing

The processing environment is the glue that must reconcile and fuse data from disparate sensors and prepare that data for rendering and display. Fusion involves incorporating both sensor data and model and simulation data such as virtual models being manipulated or data collected prior to the real-time operation in the environment. The modeling and processing environment must transform data from raw form into the form that will be presented via display interfaces to the viewer.

When data acquired from sensors exactly matches the end display device, there is little need for a modeling and processing layer. The current laparoscopic environment has been engineered so that the camera acquires an image sequence that is mapped directly onto the display device (Figure 2, left). Flexibility and power can be gained by decoupling this connection and inserting an intermediate distributed processing environment, which enables a substantial departure from this sensor-to-display direct mapping (Figure 2, right).

Since the early days of simulation technology, a number of strategies, issues, and techniques have been proposed to facilitate the development of large-scale systems to support simulations.[24] Surgical simulation environments have mirrored this trend. The Spring surgical environment,[5,6] for example, implements a multithreaded system for managing a number of sensors (haptic, camera) while running a tissue simulator based on the mass/spring physical simulator.

Toward scalable, massively parallel clusters, systems such as the Princeton Display Wall[25] use a cluster of machines to run complex simulations together with a cluster to drive the distributed rendering environment. The simulation cluster produces simulation data in real-time and the rendering cluster renders it to a scalable, large-scale display device. These parallel, distributed systems

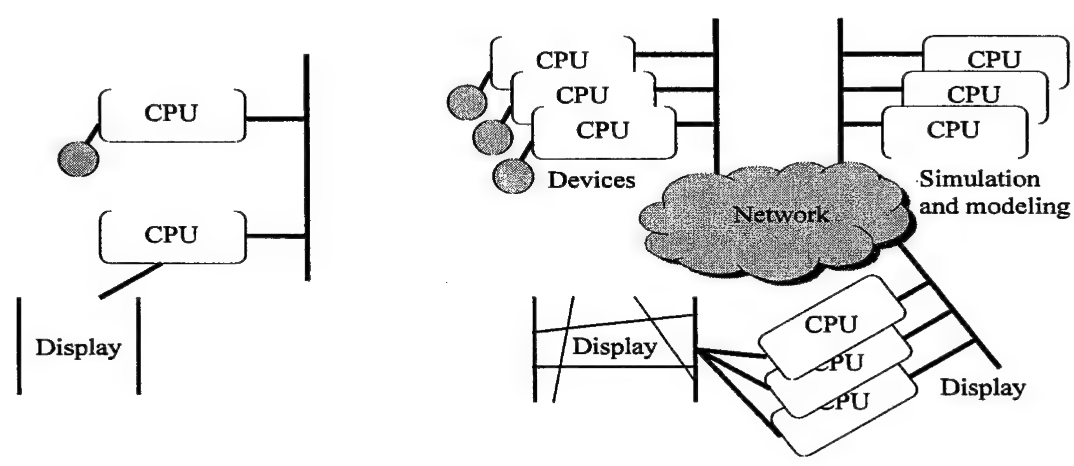

Figure 2: (left) The architecture minimizes latency by directly connecting a device (shaded circle) driven by a CPU to the display device. (right) A scalable, distributed architecture uses CPU clusters on a high-speed network to drive sets of devices, complex simulations, and a scalable, flexible display environment.

provide the required processing power to run the simulation and to drive the large-scale display.

In the case of medical simulation and more specifically laparoscopy, the stated goal[26,27] is to build a system that can reconcile imagery from the scope with preoperative imagery captured by CT scans and MRI, and any other method than might be of some service during the laparoscopic procedure. For example, comparison or recognition of anatomy with respect to a large database of procedures might be helpful. It is clear that these goals cannot be accomplished with a computing environment that so tightly couples the data acquisition devices (camera) with the display device (monitor). Parallel distributed computing environments may be the answer, and it may be the case that surgical simulation and eventually operating rooms become driven by massively parallel clusters of computers designed specifically to manage distributed sensors and preoperative, potentially collaborative patient databases.

The ultimate end of the visualization environment is the presentation of the data to the viewer, who then interprets the data and reacts, explores, and plans accordingly. Display is normally linked to the CRT, and the tight coupling that exists between the scope and the CRT reduces the flexibility that otherwise might be achieved. Visualization environments have moved toward a decoupling of individual sensors and simulation elements from the end display environment, providing flexibility in how to integrate and present information to the viewer.

Rendering and Display

Although visualization environments emphasize visible, image-driven display, they can also incorporate haptic and vestibular manipulations and audio rendering to induce spatially localized effects. High-fidelity multisensory display is desirable in order to increase bandwidth at the human–machine interface. It is implausible to expect to increase the amount of data in the system through large numbers of sensors and simulation elements while expecting the observer to continue to understand and interpret that new data without difficulty.

Along with a desire to support multisensory displays, the need for flexibility has also become apparent. The desire for flexibility in the display environment runs counter to the rigidity of most visualization tasks. Commonly available display devices, such as CRTs, stereographic systems, the Cave Automatic Virtual Environment (CAVE),[28] the Reconfigurable Advanced Visualization Environment (RAVE),[29] head-mounted displays, and other similar systems, are used to display processed data, 3D representations, and simulated environments. However, these systems are all notoriously inflexible, and many are monolithic and expensive.

How can display systems support the need for scalability, multisensory input, and flexibility? First, they leverage the distributed computing architecture (Figure 2). Second, they incorporate a special set of sensors that monitor the environment itself in order to provide flexibility. Thus, the distributed computing environment is responsible for controlling and sampling sensor input, running simulation code, controlling sensors that monitor the visualization environment itself, and combining everything into a form that can be rendered on the local display environment, which may itself be flexibly changing.

A new degree of flexibility can now be achieved through sensor monitoring. In the display realm, this has been demonstrated through camera-based approaches for geometric and photometric blending.[22,30,31] Figures 3 and 4 show how a four-projector system that is casually aligned (the projected images do not line up correctly on the display surface) can be corrected automatically. The cameras detect the geometry and compare it to what should be visible, since they can be told what to expect to see (Figure 3). Once the geometric correction is calculated, it can be applied in order to form a unified display area across the projection surface. Figure 4 (left) shows a grid after the correction is applied. The grid aligns perfectly even though the underlying projectors do not physically align.

Illumination compensation applied to areas where more than one projector overlaps makes the whole illuminated area appear to be a single image. The geometric and photometric correction, applied on-the-fly from cameras that observe how the projectors are configured, creates the appearance of a unified, seamless display that is very flexible and scalable in resolution and size. Figure 4 (right) shows an observer viewing volumetric data set projected onto a four-projector display. The contributing projectors are so well calibrated that it is very difficult to tell which projector is illuminating particular pixels on the wall.

This kind of flexibility allows projectors to be rapidly configured to cover a small amount of area at a very high number of dots per inch for applications that require close-in, very bright work. On the

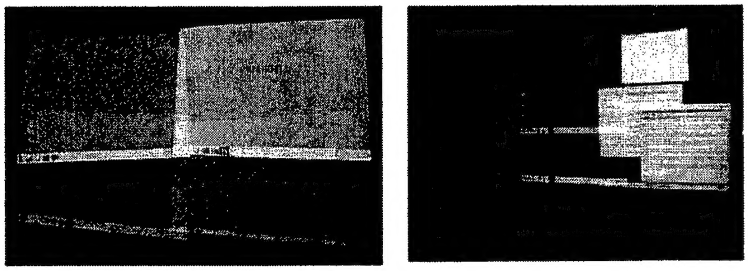

Figure 3: Projection-based display hardware can be casually aligned (left) and calibrated via automatic, camera-based algorithms (right), which calculate and apply the correct alignment transformation.

other hand, projectors can also be adjusted to cover large areas to fill entire walls with imagery. A blend is also possible: a high-resolution inset can be overlaid in an area where more detail may be required. In any case, the cameras detect the geometry and correct the imagery as required.

In the case of the surgical simulation environment, dedicated sensors can be designed to calculate the transformation between real-time scope position (perhaps given by a tracking sensor) and a set of preoperative data such as an MRI or a CT scan. Likewise, flexible, scalable display technology may help to achieve a much higher degree of flexibility and configurability for particularly unusual procedures and patient positions.

The key trends in visualization of advanced hardware capability and distributed computing environments have led to advances in data acquisition, processing and modeling, and rending and display. The distributed computing framework can incorporate a large number of independent, powerful sensors and simulation elements and helps to decouple the sensors from the end display technology, leading to flexibility and new algorithms for fusing and rendering disparate data. However, a number of issues must be addressed when these techniques are applied in the laparoscopic arena.

Visualization and Laparoscopy: Issues

Non-mission-critical visualization tasks, such as scholarly study and the exploration of digitized sculpture, or the analysis of weather patterns or fluid flow, can all be accomplished as a casual, repeatable, non-time-critical study in an environment that need not be fault tolerant. The relative lack of constraints has freed visualization researchers to explore new architectures and approaches, and much

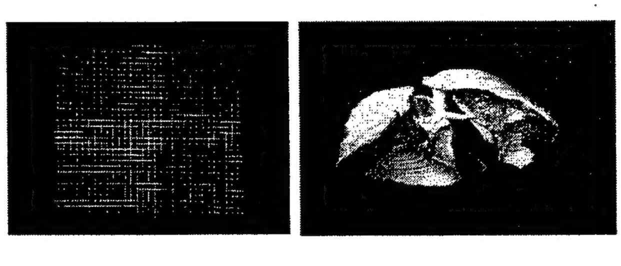

Figure 4: After calibration the projectors work cooperatively to display a seamless image.

of what has been developed is indeed intriguing. The requirements for medical analysis are much higher than the casual scientific visualization task, however. A few such requirements are:

- Equipment must be nondistracting and comfortable for the surgeon
- Latency must be minimal
- Color fidelity is crucial for identifying and assessing anatomy
- Spatial fidelity is crucial
- Calibration and registration tolerances, which can never be zero, must be very small

Equipment: Certain visualization devices are very difficult to apply in the medical realm, because they can be uncomfortable. Head-mounted displays and data gloves, for example, can be cumbersome and difficult to manage in a surgical setting. The requirement that the equipment be sterile, nondistracting, and capable of use for long periods of time without adverse effects, such as sickness and disorientation, makes it difficult to incorporate these kinds of visualization devices directly into surgical applications.

Latency: Latency is the time difference between the acquisition of data at the sensor and the display of that data to the observer. The end-to-end latency present in a system is obvious when sensor data (the image) and direct observation (cues from probes) are juxtaposed. Unfortunately, any processing between acquisition and display adds latency. The current system for laparoscopy reduces latency to a very low level by mapping the camera signal directly to the CRT. Users will likely criticize any system that creates a latency that is higher than the existing system, despite any advantages it may provide in such things as flexibility and enhancement. The latency constraint is a perfect example of how difficult it is to apply advances in visualization technology to laparoscopy. Laparoscopic systems minimize latency, and it will be difficult to move to a parallel, distributed system without sacrificing something from this constraint. One can argue that it would be easier to tolerate more latency than to work in a 3D space using only a monocular 2D image sequence. But for surgeons who are already accomplished at using the 2D, low-latency imagery to get the job done, a higher-latency system would involve loss of skill and the need for retraining.

Color fidelity: Digital cameras are built to very high standards for laparoscopy. The benefit of the high fidelity is lost, however, if this signal is displayed on a device that inadequately reproduces color. Unfortunately, many display devices do not match the high standards of the camera. In particular, commodity display devices (CRTs, LCD panels, and projectors) often trade color fidelity for a variety of other characteristics, such as brightness and portability. In particular, the commodity projectors that are used for cooperative, scalable display in visualization environments often show large variations in color and brightness characteristics.

Spatial fidelity: Mismatch between acquisition and display means that algorithms must integrate and sometimes downsample data. The algorithms for blending and fusing data can cause artifacts that mar spatial fidelity. The display has often been a limiting factor in many applications, which is why scalable display technology may be a valuable answer for expanding display devices to handle better camera resolution and to present integrated data from other devices in close proximity to camera data.

Calibration and registration: Single device calibration can often be done at the factory to specified tolerances. The distributed architecture shown in Figure 2 elevates the need for real-time, consistent, accurate calibration and registration *between* devices in a way that is, however, difficult or impossible to do off-line. Cooperative device communication and calibration is essential in the distributed environment.

Given these requirements and the high degree of reliability required of the solutions, it remains challenging to apply visualization technology in the domain of laparoscopy. Clearly, simulation environments are the first step for development, deployment, and testing. The double barrier of the operating room and the minimally invasive procedure stand as one of the most challenging environments in which to apply new visualization technology, with a potential payoff that is very high. In order to achieve a substantial breakthrough, it is crucial to embrace and exploit the trends that are driving visualization technology and to find ways to address the associated challenges that are unique to minimally invasive surgery.

References

1. Benbasat I, Todd PA: An experimental investigation of interface design alternatives: Icon vs. text and direct manipulation vs. menus. International Journal of Man-Machine Studies 38(3):369-402, 1993.

2. van Dam A: Post-WIMP user interfaces. Communications of the ACM 40(2):1997.

3. Durrani AF, Preminger GM: Three-dimensional video imaging for endoscope surgery. Comput Biol Med 25(2):237-247, 1995.

4. Sandberg W, Ganous T, Steiner C: Setting a research agenda for perioperative systems design. Semin Laparosc Surg 10(2):57-70.

5. Montgomery K, Bruyns C, Brown J, Sorkin S: Spring: A general framework for collaborative, real-time surgical simulation in: Westwood J et al (eds.): Medicine Meets Virtual Reality; Amsterdam; IOS Press, 2002.

6. Montgomery K, Heinrichs L, Bruyns C, Wildermuth S, et al: Surgical simulator for hysteroscopy: A case study of visualization in surgical training. Presented at IEEE Visualization 2001, San Diego Paradise Point Resort, San Diego, October 21-26, 2001.

7. McCormick BH, DeFanti: Special report: Visualization in scientific computing—a synopsis. IEEE Computer Graphics and Applications (July 1987), pp 61-70.

8. http://www.strykercorp.com/. Last accessed August 6, 2003.

9. http://www.intuitivesurgical.com/. Last accessed August 6, 2003.

10. Rattner D, Park A: Advanced devices for the Operating Room of the Future. Semin Laparosc Surg 10(2):85-89.

11. Moore G: Nanometers and gigabucks--Moore on Moore's Law. Distinguished Lecture, 1996.

12. Lundstrom M: Is nanoelectronics the future of microelectronics? In: Source International Symposium on Low Power Electronics and Design. Monterey, Calif, 2002, pp 172-177.

13. Fuchs H, Livingston MA, Raskar R, Colucci D, et al: Augmented reality visualization for laparoscopic surgery. Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention, pp 934-943.

14. http://www.inneroptic.com. Last accessed August 6, 2003.

15. http://smartcamera.stanford.edu. Last accessed August 6, 2003.

16. Yang R, Brown MS, Seales WB, Fuchs H: Geometrically correct imagery for teleconferencing. In: Proceedings of the Seventh ACM International Conference on Multimedia, October 30-November 5, 1999.

17. Crockett TW, Orloff T: A MIMD rendering algorithm for distributed memory architectures. 1993 Parallel Rendering Symposium Proc; IEEE, October 1993, pp 35-42.

18. Humphreys G, Hanrahan P: A distributed graphics system for large tiled displays. Proceedings of Visualization '99, IEEE, October 1999, pp 215-223.

19. Muraki S, Ogata M, Ma K-L, Koshizuka K, et al: Next-generation visual supercomputing using PC clusters with volume graphics hardware devices. ACM/IEEE Conference on Supercomputing, November 2001.

20. Rhomberg A, Enzler R, Thaler M, Tröster G: Design of a FEM computation engine for real-time laparoscopic surgery simulation. Proceedings of the First Merged International Parallel Processing Symposium and Symposium on Parallel and Distributed Processing, Los Alamitos, California, 1998, pp 711-715

21. Guthart G Salisbury J Jr: The intuitive telesurgery system: Overview and application. Proc. IEEE ICRA, 2000.

22. Jaynes C, Seales B, Calvert K, Fei Z, et al: The Metaverse - a collection of inexpensive, self-configuring, immersive environments. 7th International Workshop on Immersive Projection Technology/Eurographics Workshop on Virtual Environments, Zurich Switzerland, May 2003.

23. Chandrakasan A, Min R, Bhardwaj M, Cho S-H, et al: Power aware wireless microsensor systems. Keynote paper, ESSCIRC, Florence, Italy, September 2002.

24. Clema JK: General purpose tools for system simulation. 11th Annual Simulation Symposium Tampa, Florida, 1978, pp 37-60.

25. Chen H, Wallace G, Gupta A, Li K, et al: Experiences with scalability of display walls. In: Immersive Projection Technology Symposium (IPT), March 2002.

26. http://www.orofthefuture.org/overview/. Last accessed August 6, 2003.

27. Grimson E, Leventon M, Ettinger G, Chabrerie A, et al: Clinical experience with a high precision image-guided neurosurgery system. MICCAI, 1998, pp 63-73.

28. Cruz-Neira C, Sandin D, DeFanti T: Virtual reality: The design and implementation of the CAVE. Proceedings of SIGGRAPH 93 Computer Graphics Conference, ACM SIGGRAPH, August 1993, pp 135-142.

29. http://www.fakespacesystems.com/. Last accessed August 6, 2003.

30. Raskar R, Brown M, Yang R, Chen W, et al: Multi-projector displays using camera-based registration. IEEE Visualization '99, 1999.

31. Brown M, Seales W, Webb S, Jaynes C: Building large format displays for digital libraries. Communications of the ACM 44(5): 2001.